

8D Likelihood Effective Higgs Couplings Extraction Framework in  $h \rightarrow 4\ell$ 

Yi Chen,<sup>\*</sup> Emanuele Di Marco,<sup>†</sup> and Maria Spiropulu<sup>‡</sup>  
*Lauritsen Laboratory of Physics, California Institute of Technology,  
 1200 E. California Blvd, Pasadena, CA, 92115, USA*

Joe Lykken<sup>§</sup>  
*Theoretical Physics Department, Fermilab, P.O. Box 500, Batavia, IL 60510, USA*

Roberto Vega-Morales<sup>¶</sup>  
*Laboratoire de Physique Théorique d'Orsay, UMR8627-CNRS,  
 Université Paris-Sud 11, F-91405 Orsay Cedex, France  
 Theoretical Physics Department, Fermilab, P.O. Box 500, Batavia, IL 60510, USA and  
 Department of Physics and Astronomy, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA*

Si Xie<sup>\*\*</sup>  
*Lauritsen Laboratory of Physics, California Institute of Technology, Pasadena, CA, 92115*

We present an overview of a comprehensive analysis framework aimed at performing direct extraction of all possible effective Higgs couplings to neutral electroweak gauge bosons in the decay to electrons and muons, the so called ‘golden channel’. Our framework is based primarily on a maximum likelihood method constructed from analytic expressions of the fully differential cross sections for  $h \rightarrow 4\ell$  and for the dominant irreducible  $q\bar{q} \rightarrow 4\ell$  background, where  $4\ell = 2e2\mu, 4e, 4\mu$ . Detector effects are included by an explicit convolution of these analytic expressions with the appropriate transfer function over all center of mass variables. Utilizing the full set of observables, we construct an un-binned detector-level likelihood which is continuous in the effective couplings. We consider possible  $ZZ$ ,  $Z\gamma$ , and  $\gamma\gamma$  couplings simultaneously, allowing for general CP odd/even admixtures. A broad overview is given of how the convolution is performed and we discuss the principles and theoretical basis of the framework. This framework can be used in a variety of ways to study Higgs couplings in the golden channel using data obtained at the LHC and other future colliders.

## I. INTRODUCTION

The recent discovery of the Higgs boson at the LHC [1, 2] with properties resembling those predicted by the Standard Model, shifts our attention to the determination of its precise nature and to establish whether or not the Higgs boson possesses any anomalous couplings to Standard Model particles. In this study we focus on couplings to neutral electroweak gauge bosons. Since these ‘anomalous effects’ are expected to be small if at all present, constraining or measuring of these couplings should preferably be done through direct parameter extraction with minimal theoretical assumptions. The vast literature [3–37] on Higgs decays to four charged leptons (electrons and muons) through neutral electroweak gauge bosons, suggests that the so called ‘golden channel’, can be a powerful means towards accomplishing this goal.

A number of frameworks have been established utilizing the Matrix Element Method to study the golden channel aiming to determine these potentially anomalous

couplings. These primarily rely on Monte Carlo generators such as the JHU generator [13, 17, 32] or on Madgraph implementations [22, 31]. They have the advantage of flexibility to include various Higgs production and decay channels and are especially useful for constructing kinematic discriminators to distinguish between competing hypotheses.

Focusing on the golden channel only<sup>1</sup>, we propose a novel analysis framework largely based on an analytic implementation. It is designed to maximize the information contained in each event with the aim of direct extraction of the various effective Higgs couplings. It is generally acknowledged in the literature that analytic methods are optimal for performing this direct multi-parameter extraction within practical and reasonable computational processing resources [13, 17, 32]. In this work, we also demonstrate that within an analytic framework one can readily include the relevant detector effects and obtain a detector-level likelihood function in terms of the full set of observables available in the four lepton final state. This is accomplished by the explicit convolution of analytic expressions for the ‘truth level’ fully differential cross sections with a transfer function which parametrizes the detector resolution and acceptance effects.

<sup>\*</sup> yi.chen@cern.ch  
<sup>†</sup> emanuele.di.marco@cern.ch  
<sup>‡</sup> smaria@caltech.edu  
<sup>§</sup> lykken@fnal.gov  
<sup>¶</sup> roberto.vega@th.u-psud.fr  
<sup>\*\*</sup> sixie@hep.caltech.edu

<sup>1</sup> Though we will not discuss it explicitly here, we are also able to extend our framework to the  $h \rightarrow \gamma\gamma$  and  $h \rightarrow 2\ell\gamma$  channels.

This analysis framework has already proved useful in constraining effective Higgs couplings as demonstrated in a recent CMS analysis [38, 39]. It was shown that for simplified cases of constraining one or two parameters, our framework gives comparable performance to other established analysis methods for the golden channel [13, 17, 22, 31, 32, 38, 39]. In this work we present an overview of the framework and discuss the principles and theoretical basis. In particular we sketch how the various components of the detector level likelihood are constructed with emphasis on how the convolution integral is performed as well as various validations. How the likelihood can then be used to perform multi-parameter extraction of effective Higgs couplings is also discussed. We hope that the additional features of our framework are also found useful in the next phase of the LHC and future colliders. Much more information on the framework including technical details can be found in [19, 35, 37–41].

## II. OVERVIEW OF FRAMEWORK

Though ‘truth’ level (or generator) studies of  $h \rightarrow 4\ell$  ( $4\ell = 2e2\mu, 4e, 4\mu$ ) give a good approximate estimate of the expected sensitivity to the Higgs  $ZZ$ ,  $Z\gamma$ , and  $\gamma\gamma$  couplings [37], when analyzing data obtained at the LHC (or future colliders) a detector level likelihood which accounts for the various detector effects is necessary. Since generally detector level likelihoods are obtained via the use of Monte Carlo methods, it becomes difficult to obtain the full multi-dimensional likelihood for the  $4\ell$  final state. Typically one needs to fill large multi-dimensional templates that require an impractical amount of computing time. There are also potential collateral binning and ‘smoothing’ side-effects often associated with these methods. In the case of the golden channel this necessitates the use of kinematic discriminants which ‘collapse’ the fully multi-dimensional likelihood into two or perhaps three detector level observables [32]. This approach is normally taken to facilitate the inclusion of detector effects, but is not optimal when fitting to a large number of parameters simultaneously [17, 42]. This is unfortunate in the case of the golden channel where in principle there are twelve observables which can be used to extract a large number of parameters at once, including their correlations. It would be satisfying and useful to have a framework which is free of these issues and capable of utilizing all available information in the four lepton final state at detector level.

### A. From ‘Truth’ to ‘Detector’ Level

This is accomplished in our framework by performing an explicit convolution of the generator (‘truth’) level probability density, formed out of the signal and background differential cross sections, with a transfer function which encapsulates the relevant detector effects. This can

be represented schematically as follows,

$$P(\vec{X}^R|\vec{\mathcal{A}}) = \int P(\vec{X}^G|\vec{\mathcal{A}})T(\vec{X}^R|\vec{X}^G)d\vec{X}^G. \quad (1)$$

Here we take  $\vec{X}$  to represent the full set of center of mass variables, of which there are twelve in the golden channel, to be discussed more below, and  $\vec{\mathcal{A}}$  represents some set of lagrangian parameters. The transfer function  $T(\vec{X}^R|\vec{X}^G)$  takes us from generator (G) level to reconstructed (R) level observables and represents the probability of reconstructing the observables  $\vec{X}^R$  given the generator level observable  $\vec{X}^G$ . It is treated as a function of  $\vec{X}^R$  which takes  $\vec{X}^G$  as input. As will be described more in Sec. V A, once the integration in Eq.(1) is performed we must then normalize over all twelve reconstructed level observables to obtain the detector level *pdf*.

The integral in Eq.(1) is the defining feature of our framework and has been obtained for both the  $h \rightarrow 4\ell$  signal as well as the dominant  $q\bar{q} \rightarrow 4\ell$  background, which have been computed analytically in accompanying studies [19, 35, 43]. We emphasize that the integral has not been obtained via Monte Carlo methods. Instead we have explicitly performed the integration by utilizing various analytic and well-established numerical methods [41, 44] (for studies that perform similar convolutions using Monte Carlo methods see [31, 45–47]). This ensures that (arbitrarily) high precision is maintained at each step, producing what is effectively an ‘analytic function’ in terms of detector level variables once the convolution has been performed. After performing this 12-dimensional integration and normalizing, we are left with a probability density function (*pdf*) from which we *construct an un-binned twelve-dimensional detector level likelihood* which is a continuous function of the effective couplings (or Lagrangian parameters) and takes as its input, up to twelve reconstructed (detector-level) center of mass observables. In the current implementation we will average over the four production variables to reduce the systematic uncertainties, thus obtaining an eight-dimensional likelihood in terms of just decay observables. However, this step is in principle not necessary.

We also emphasize that the convolution integral is largely independent of detector transfer function and generator level differential cross section and in particular how accurate the descriptions of the ‘truth’ level *pdfs* or the detector properties are. This means the framework can in principle be adapted to any detector which studies  $h \rightarrow 4\ell$  (or any  $X \rightarrow 4\ell$ ) and, in addition, as theoretical calculations of the generator level differential cross sections improve they can easily be incorporated into the convolution integral. Thus, there is ample room for optimization in our framework as time goes on. The generality of the convolution also allows for other beyond the Standard Model physics such as exotic Higgs decays [48] to be easily be incorporated into the  $h \rightarrow 4\ell$  framework.

## B. Analytic Parameterizations

As we discuss below, it is essential to first have analytic parameterizations of the ‘truth’ level differential cross sections in order to perform the convolution integral in Eq.(1). These can be obtained in essentially two different ways. The first is to simply analytically compute the differential cross section starting from Feynman diagrams, which of course is not always possible. The second is to obtain an ‘analytic’ parameterization by fitting to a large Monte Carlo sample with some appropriately parametrized function. This becomes quite difficult when the function is multi-dimensional as is the case in the golden channel with twelve center of mass observables and requires large samples and an accurate interpolation procedure. In this framework we have implemented a hybrid of these two approaches with the primary component coming from analytic expressions of the leading order  $h \rightarrow 4\ell$  and  $q\bar{q} \rightarrow 4\ell$  fully differential cross sections [19, 35, 43].

Since NLO effects in the golden channel are generally small [49–51], these leading order differential cross sections represent the dominant contributions to the  $4\ell$  ‘truth’ level likelihood. There are however, a number of sub-dominant effects which appear at higher order and should be accounted for. These include production and additional background effects. In these cases, the second method of parametric fits to simulated data is typically the optimal route. To do this we follow a similar procedure as found in [32] while further details on the implementation into our framework can be found in [41]. We emphasize however that the convolution integral is independent of these matters allowing for easy implementation of more precise ‘truth’ level likelihoods as they become available over time.

Of course when considering the detector level likelihood there are additional, but again sub-dominant, effects not present at ‘truth’ level which should be accounted for, such as detector momentum resolution and acceptance effects. These can be parametrized via transfer functions which can be optimized for a particular detector as done recently in [38, 39] which incorporates a parameterization of the CMS detector into our framework. Since these typically would be supplied by the experimentalist we do not discuss their construction in detail here, but note that as knowledge of the detectors improves and parameterizations of the transfer functions become more accurate, they can easily be incorporated into the convolution integral in Eq.(1), but again the integration is independent of these matters. More details on the construction and implementation, as well as the validation, of the transfer functions is found in [38, 39, 41].

## C. Fast Parameter Extraction

The convolution integral in Eq.(1) allows us to (effectively) obtain an analytic function in both detector

level observables and lagrangian parameters. This is because, via the explicit 12-dimensional integration over all center of mass variables, we are able to obtain a 1-to-1 mapping from the ‘truth’ level likelihood to the ‘detector’ level likelihood. This allows us, during parameter extraction, to effectively work directly with the lagrangian parameters, but at detector level which gives us the ability to easily perform multi-parameter extraction with the same speed and flexibility as was done at generator level [35, 37]. Being able to fit to multiple parameters simultaneously is important since it allows for strong tests of models which often predict correlations between the various parameters. We point out that our framework allows us to do this while avoiding relying on hypothesis testing or on the construction of kinematic discriminants which is less optimal when extracting multiple parameters than maximizing the full likelihood [52] where all observables are used. Furthermore, the analytic nature of our framework allows for a great deal of flexibility in performing a variety of types of parameter extractions and re-parameterizations.

## D. Comments on Assumptions and Approximations

In performing the convolution integral in Eq.(1) we have relied on two key assumptions. The first is that angular resolution effects due to detector smearing can be neglected, which is an excellent approximation for the LHC detectors [53–55]. Second, we have assumed in the transfer function that each lepton is independent of the others which again is a very good approximation since leptons are clean and well-measured objects in the CMS and ATLAS detectors once standard lepton selection criteria are imposed [53–55]. With these simplifying assumptions the convolution integral can then be performed as will be described below and in much more detail in [40] and [41].

Even after the convolution is performed however, we must still normalize the detector level differential cross section. Since this can not be done analytically one must resort to Monte Carlo techniques. Thus, strictly speaking the final *pdf* is not analytic. However, as we will discuss more in Sec. V A, due to the manner in which the analytic expressions are organized, a high precision on the normalization can be obtained in a short amount of computing time. Furthermore, by fitting to ratios of couplings, we can circumvent the need for the *absolute* normalization which greatly simplifies the computational procedure and allows us to achieve a high precision [40, 41] leading in the end to a detector level *pdf* which is *effectively analytic* in reconstructed observables and lagrangian parameters.

Of course there are components of both the ‘truth’ and detector level likelihoods which can not be included in the convolution integral of Eq.(1). These correspond to any components for which a sufficiently accurate analytic parameterization can not be obtained. These may

include potential higher order contributions to both signal and background differential cross sections as well as additional fake backgrounds such as  $Z + X$ . For these one must resort to more conventional Monte Carlo techniques and the construction of large (binned) ‘look-up’ tables. The effects of binning can be mitigated through a linear multi-dimensional interpolation technique which is described in more detail in [41]. Fortunately these components are sub dominant in the golden channel and can be assigned systematics to study their effects [38, 39, 41].

There are also various other systematics associated with both detector and theoretical uncertainties which should properly be accounted for. Since these components do not have a large effect on the final sensitivity (especially once sizable data sets are accumulated) and are not directly related to the convolution, we will discuss them only briefly below, but see [38, 39, 41] for more details on how they are implemented into the framework in a real experimental analysis.

In constructing the detector level likelihood we have overcome many of the technical challenges which in the past have made it impossible to use the fully multi-dimensional likelihood during parameter fitting. Below we sketch in more detail how these various challenges have been overcome, but many of the details are technically beyond the scope of this paper so we refer the reader to [19, 35, 37–41] for more details.

### III. THE ‘TRUTH LEVEL’ PDF

Before obtaining the detector level likelihood one must of course first construct the ‘truth’ level (or generator level) likelihood. As we discuss, the generator level likelihood is composed of a ‘decay’ and ‘production’ differential spectrum. In our framework, the primary component is constructed out of analytic expressions for the  $h \rightarrow 4\ell$  signal and the dominant  $q\bar{q} \rightarrow 4\ell$  background differential cross sections. Analytic expressions have been shown to be useful in likelihood methods where the full kinematics of an event can be exploited. This is especially true for the golden channel as has been demonstrated in numerous studies [13–15, 17, 18, 29, 32, 35, 37]. For a detailed description of the analytic calculations for the signal and background fully differential cross sections as well as their validation we refer the reader to accompanying studies [19, 35].

Below we give an overview of how the ‘truth’ level likelihood is constructed and define the twelve center of mass variables in the four lepton final state. We briefly discuss our parameterization of the Higgs couplings to electroweak gauge bosons and how the analytic expressions are combined with the appropriate production spectra to form the full truth level differential cross section. We also discuss in this section how the production spectrum is obtained and comment on the additional backgrounds present in the golden channel.

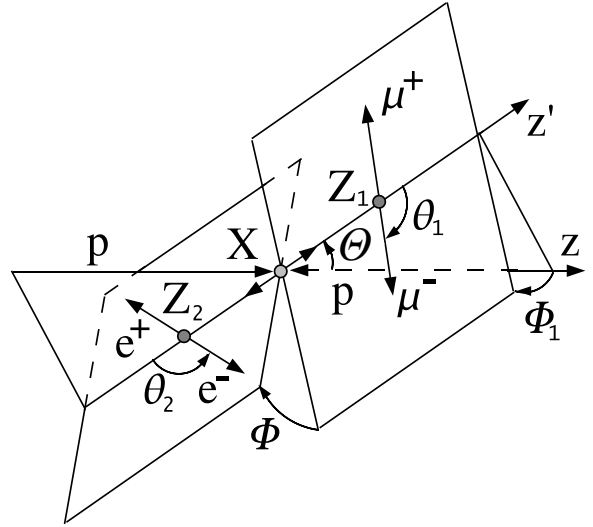


FIG. 1. Definition of angles in the four lepton center of mass frame  $X$ .

#### A. Center of Mass Observables

Here we describe the various center of mass variables which will be used as our set of observables when constructing the likelihood. The kinematics of four lepton events are illustrated in Fig. 1. The invariant masses are defined as the following:

- $\sqrt{\hat{s}} \equiv M_{4\ell} \equiv m_h$  – The invariant mass of the four lepton system or the Higgs mass in case of signal.
- $M_1$  – The invariant mass of the lepton pair system which reconstructs closest to the  $Z$  mass.
- $M_2$  – The invariant mass of the other lepton pair system and interpreted as  $M_2 < M_1$ . This condition holds as long as  $\sqrt{\hat{s}} \gtrsim 2m_Z$ .

These invariant masses are all independent subject to the constraint  $(M_1 + M_2) \leq \sqrt{\hat{s}}$  and serve as the most strongly discriminating observables between different signal hypothesis as well as between signal and background. Note also that the  $4e/4\mu$  final state can be reconstructed in two different ways due to the identical final state interference. This is a quantum mechanical effect that occurs at the amplitude level and thus both reconstructions are valid. The definitions  $M_1$  and  $M_2$  remain unchanged however.

The angular variables are defined as:

- $\Theta$  – The production angle between the momentum vectors of the lepton pair which reconstructs to  $M_1$  and the total  $4\ell$  system momentum.
- $\theta_{1,2}$  – Polar angle of the momentum vectors of  $e^-, \mu^-$  in the lepton pair rest frame.
- $\Phi_1$  – The angle between the plane formed by the  $M_1$  lepton pair and the ‘production plane’ formed

out of the momenta of the incoming partons and the momenta of the two lepton pair systems.

- $\Phi$  – The angle between the decay planes of the final state lepton pairs in the rest frame of the  $4\ell$  system.

We group the angular variables as follows  $\vec{\Omega} = (\Theta, \cos \theta_1, \cos \theta_2, \Phi_1, \Phi)$ . These angular variables are useful in aiding to distinguish different signal hypothesis and in particular between those with different CP properties, as well as in discriminating signal from background. There are also additional production variables associated with the initial partonic state four momentum:

- $\vec{p}_T$  – The momentum in the transverse direction.
- $Y$  – Defined as the motion along the longitudinal direction.
- $\phi$  – Defines a global rotation of the event in the  $4\ell$  rest frame and in general does not aid greatly in discriminating power.

Including  $Y$  and  $\vec{p}_T$  as observables in the likelihood increases the discriminating power of the golden channel. However, including these production variables can introduce large uncertainties since their spectra includes

parton distribution functions as well as NLO contributions which should be included. Thus, they are often integrated out of the final likelihood [38, 39] and not used during parameter extraction. This reduces the potential discriminating power, but this is compensated by the smaller systematic uncertainties one obtains by not including them in the likelihood (or averaging over them). We will discuss more below how in our framework one can easily either include them in the final likelihood or average over them to mitigate the effects of the uncertainties associated with these variables. However as with  $\phi$ , it is crucial to include them when performing the convolution with the transfer function in order to obtain the proper detector level likelihood. These variables exhaust the twelve possible center of mass observables available in the golden channel.

## B. Parameterization of Scalar-Tensor Couplings

Assuming only Lorentz invariance, the general couplings of a spin-0 particle to two spin-1 vector bosons can be parametrized in terms of effective couplings by the following tensor structure,

$$\Gamma_i^{\mu\nu} = \frac{i}{v} \left( \mathcal{A}_1^i m_Z^2 g^{\mu\nu} + \mathcal{A}_2^i (k_1^\nu k_2^\mu - k_1 \cdot k_2 g^{\mu\nu}) + \mathcal{A}_3^i \epsilon^{\mu\nu\alpha\beta} k_{1\alpha} k_{2\beta} + \left( \mathcal{A}_4^i \left( \frac{k_1^2 + k_2^2}{m_Z^2} \right) + \mathcal{A}_5^i \left( \frac{\hat{s}}{m_h^2} \right) \right) m_Z^2 g^{\mu\nu} \right), \quad (2)$$

where in the golden channel  $i = ZZ, Z\gamma, \gamma\gamma$ . The variables  $k_1$  and  $k_2$  represent the four momentum of the intermediate vector bosons with  $v$  the Higgs vacuum expectation value (vev) which we have chosen as our overall normalization. The  $\mathcal{A}_n^i$  are dimensionless and in principle arbitrary complex form factors with possible momentum dependence (or more precisely a  $\hat{s}, k_1^2, k_2^2$  dependence) making Eq.(2) completely general. Note that the tensor structure for  $\mathcal{A}_5^i$  is only distinguishable from  $\mathcal{A}_1^i$  for

off-shell Higgs decays as discussed in [36]. For a purely Standard Model Higgs we have  $\mathcal{A}_1^{ZZ} = 2$  at tree level while all other effective couplings are generated at higher loop order and at most  $\mathcal{O}(\lesssim 10^{-2})$ .

Of course it is often possible to expand the  $\mathcal{A}_n^i$  in a power series of momenta keeping only the leading (constant) terms. By keeping the leading terms in this expansion there is a one-to-one mapping from this vertex onto the effective Lagrangian<sup>2</sup>,

$$\mathcal{L} \supset \frac{1}{4v} \left( 2\mathcal{A}_{1o}^{ZZ} m_Z^2 h Z^\mu Z_\mu + \mathcal{A}_{2o}^{ZZ} h Z^{\mu\nu} Z_{\mu\nu} + \mathcal{A}_{3o}^{ZZ} h Z^{\mu\nu} \tilde{Z}_{\mu\nu} - 4\mathcal{A}_{4o}^{ZZ} h Z_\mu \square Z^\mu - 2\mathcal{A}_{5o}^{ZZ} \left( \frac{m_Z}{m_h} \right)^2 \square h Z_\mu Z^\mu \right. \\ \left. + 2\mathcal{A}_{2o}^{Z\gamma} h F^{\mu\nu} Z_{\mu\nu} + 2\mathcal{A}_{3o}^{Z\gamma} h F^{\mu\nu} \tilde{Z}_{\mu\nu} + \mathcal{A}_{2o}^{\gamma\gamma} h F^{\mu\nu} F_{\mu\nu} + \mathcal{A}_{3o}^{\gamma\gamma} h F^{\mu\nu} \tilde{F}_{\mu\nu} \right), \quad (3)$$

where the  $\mathcal{A}_{no}^i$  represent the leading, momentum independent coefficients and electromagnetic gauge invariance requires  $\mathcal{A}_{1o}^{Z\gamma, \gamma\gamma} = \mathcal{A}_{4o}^{Z\gamma, \gamma\gamma} = \mathcal{A}_{5o}^{Z\gamma, \gamma\gamma} = 0$ . We have

defined  $Z^\mu$  and  $A^\mu$  as the  $Z$  and photon fields respectively while  $V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu$  are the usual bosonic field strengths and the dual field strengths are defined as  $\tilde{V}_{\mu\nu} = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} V^{\rho\sigma}$ . The effective lagrangian in Eq.(3) is composed of the leading terms in a derivative expansion (up to two derivatives) and is useful for parametrizing potentially large new physics effects generated by loops

<sup>2</sup> This lagrangian has been implemented [56] into the FeynRules/Madgraph [57, 58] framework for validation purposes.

of heavy particles and a convenient framework for assessing the potential sensitivity to the leading operators [37] involving photons and  $Z$  bosons.

Thus, although Eq.(2) is a redundant parameterization of the tensor structure, it is a convenient, yet more general, parametrization for fitting to effective Lagrangian parameters that might be generated in various models at dimension five or less as in Eq.(3). The parameterization in Eq.(2) can of course be mapped, with appropriate translation of the parameters, onto Lagrangians with dimension greater than five or to an underlying dimension six lagrangian in a theory of electroweak symmetry breaking such as in the Standard Model. We will work explicitly with the vertex in Eq.(2) which has been used to calculate the fully differential cross section for  $h \rightarrow 4\ell$  and when performing parameter extraction, but other parameterizations can be easily accommodated.

This flexibility of parameterization also allows for other new physics, such as exotic Higgs decays involving exotic fermions or vector bosons [48], to be easily included in the framework. Furthermore, by using this parameterization, explicit computations of either Standard Model or new physics loop effects which would generate these momentum dependent form factors can easily be included into the framework. This allows for the ability to in principle extract the parameters from whichever underlying theory is responsible for generating them. We leave a more detailed investigation of these loop effects to ongoing work [59].

### C. Signal and Background Fully Differential Cross Sections

In the case of signal we have computed analytically the fully differential cross section in the observables described in Sec.III A for the process  $h \rightarrow ZZ + Z\gamma + \gamma\gamma \rightarrow 4\ell$  using the parameterization in Eq.(2). We have included all possible interference effects between tensor structures as well as identical final states in the case of  $4e/4\mu$ . For the irreducible background we have computed analytically the process  $q\bar{q} \rightarrow ZZ + Z\gamma + \gamma\gamma \rightarrow 4\ell$  which includes the s-channel (resonant)  $4\ell$  process as well as the t-channel (diboson production)  $4\ell$  process and again have included all possible interference effects. All vector bosons are allowed to be on or off-shell and we do not distinguish between them in what follows. The details of these calculations can be found in [19, 35, 37, 43] along with the validation procedures and studies of the distributions as well as the various interference effects. We have combined these analytic expressions with functions parametrizing the production spectra and implemented them into our analysis framework.

We note that it is important to include all possible Higgs couplings including the  $Z\gamma$  and  $\gamma\gamma$  contributions in the signal differential cross section since the Higgs appears to be mostly Standard Model-like [60] and we are primarily searching for small anomalous deviations from

the Standard Model prediction. Thus when attempting to extract specific couplings we must be sure that one small effect is not being mistaken for another. This is particularly relevant since many of the couplings may be correlated with one another.

Furthermore, it has been shown recently [37] that for ‘true’ points near the Standard Model, the greatest sensitivity to the anomalous couplings (non  $\mathcal{A}_{10}^{ZZ}$ ) is for the  $Z\gamma$  and especially  $\gamma\gamma$  operators (see Eq.(3)). Including all possible couplings and doing a simultaneous fit ensures that we minimize the possibility of misinterpretation or of introducing a bias when attempting to extract these couplings. Searching for these small effects is also why it is important to include the interference effects between the identical final state leptons as well as the relevant detector effects and background.

We also comment that in principal there are NLO contributions to the  $h \rightarrow 4\ell$  decay processes, but these are expected to be small at  $\sim 125$  GeV [49, 50] and not relevant until higher precision is obtained once larger data sets are gathered. Eventually however, these effects should be included and their implementation into our framework is currently ongoing.

### D. Combining Production and Decay

To be able to perform a fit for the effective Higgs couplings, we must first construct the fully differential cross section for the observables as a function of the undetermined parameters ( $\vec{A}$ ). This differential cross section consists of two components which we assume to be factorized: the parton level (‘decay’) differential cross section as discussed in Section III C, and the ‘production’ spectrum. The full production plus decay fully differential cross section can be expressed as the following,

$$P(\vec{p}_T, Y, \phi, \hat{s}, M_1, M_2, \vec{\Omega}|\vec{A}) = \frac{dW_{\text{prod}}(\hat{s}, \vec{p}_T, Y, \phi)}{d\hat{s}dYd\vec{p}_Td\phi} \times \frac{d\sigma_{4\ell}(\hat{s}, M_1, M_2, \vec{\Omega}|\vec{A})}{dM_1^2 dM_2^2 d\vec{\Omega}}, \quad (4)$$

where, since the Higgs is a spin-0 particle, we can explicitly assume that the decay process can be factorized from the production mechanism. For the background this explicit factorization does not occur, but still turns out to be an adequate approximation [41] especially if the  $\vec{p}_T$  and  $Y$  variables are averaged over once the convolution is performed. The parton level fully differential cross section ( $\sigma_{4\ell}$ ) is treated as being at fixed  $\hat{s}$  where one obtains the input  $\hat{s}$  value from the production spectrum ( $W_{\text{prod}}$ ). The production spectrum for the signal and background depend on the parton distribution functions and can not be computed analytically. For the signal which we assume decays on-shell, the  $\hat{s}$  spectrum is taken to be a delta function centered at  $m_h^2$ , which for a Standard Model Higgs at 125 GeV is an excellent approximation. Note however that this assumption can be relaxed in our framework to consider more general  $\hat{s}$  spec-

tra as would be found for example in the case of a new heavy scalar with a large width.

### E. Comments Production Spectra

Here we discuss how the  $W_{\text{prod}}(\hat{s}, \vec{p}_T, Y, \phi)$  production spectrum in Eq.(4) is obtained. This function involves higher order effects as well as parton distribution functions and thus can not be computed analytically. To include them in the total differential cross sections there are various options. One can in principal generate enough Monte Carlo events to accurately fill the full spectrum in the  $(\hat{s}, Y, \vec{p}_T, \phi)$  variables. As this is computationally intensive we take an approximate approach in which we interpolate analytic functions for the signal and background from one or two dimensional projections generated from the Madgraph [61] and POWHEG [62] Monte Carlo generators following a similar procedure as found in [32]. Having an analytic parameterization for these functions also allows for faster integration when implementing them into the convolution procedure described above. This procedure of interpolating the one or two dimensional projections neglects correlations between the production variables. However, since in the signal case there is an explicit factorization between production and decay, the effects of this approximation on parameter extraction in our analysis are small.

In addition, to mitigate these effects further, one can always average over  $Y$  and  $\vec{p}_T$  as well as fit to ratios of couplings while taking the Higgs mass and overall normalization as input from the total rate (so called ‘geolocating’ [36]) as was done in a recent implementation of our framework into a CMS experimental analysis [38, 39]. Note however, the overall normalization and Higgs mass can in principle be extracted in our framework, but as this requires extra careful treatment of the production spectra and additional backgrounds we defer a discussion of this to future work.

### F. Comments on Additional Backgrounds

For the background there are also the higher order contributions such as the  $gg \rightarrow 4\ell$  and  $Z + X$  processes. These make up the parts of the likelihood which can not currently be included in the convolution integral since a sufficiently accurate analytic parameterization has yet to be obtained. Thus for these components we must resort to constructing large ‘look-up’ tables via Monte Carlo generation. Again, the effects of the necessary binning can be mitigated through a linear multi-dimensional interpolation technique [41]. Additionally, there will be systematic uncertainties associated with these components. Fortunately, the  $gg \rightarrow 4\ell$  component only makes up  $\sim 3 - 5\%$  relative to  $q\bar{q} \rightarrow 4\ell$  around 125 GeV [51]. The  $Z + X$  background on the other hand does make up a sizable contribution of the

total background which is comparable to, but smaller than, the largest  $q\bar{q}$  component (see Table 2 in [38] or Table 3 in [39]). This component however can in principle be reduced further in the future by requiring more stringent lepton acceptance criteria once more data is collected. For now the use of the linearly interpolated ‘look-up’ templates and associated systematics is found to be sufficient. Once the templates are built, including these components in the final likelihood is straightforward as we briefly sketch in Sec. VB and discussed in more detail in [38, 39, 41].

## IV. THE ‘DETECTOR LEVEL’ PDF

The convolution integral in Eq.(1) is conceptually straightforward, but in practice is challenging to perform, both for computational and algorithmic reasons. The key assumption which makes it possible is that the *direction* of lepton momenta are measured with infinite precision which at CMS and ATLAS is a very good approximation. This allows us, through a change of variables, to reduce the 12-dimensional integral into a more manageable 4-dimensional integral over the four energies of the leptons which are altered by detector resolution effects. Typically this 4-dimensional integral is done using Monte Carlo techniques [31], thus losing the advantage of having analytic control over the likelihood or assuming that the resolution effects can also be neglecting making the integral trivial. As discussed in Sec. II we instead perform this integration explicitly using a combination of numerical and analytic methods which allow us to maintain arbitrarily high precision at each step involved. There are a number of technical details involved in this procedure which are beyond the scope of this ‘overview’ of the framework, but the details can be found in [40, 41]. We instead briefly sketch an overview of the convolution integral and show its validation.

### A. Transforming from CM Basis to Lepton Smearing Basis

Beginning from Eq.(1) we first discuss the construction of the background detector level *pdf*. The construction of the signal will be discussed separately as there is a subtle, but important, difference in performing the convolution. Since there are no undetermined parameters in the background the generator and detector-level (unnormalized) differential cross sections are given simply by  $P_B(\vec{X}^G)$  and  $P_B(\vec{X}^R)$  respectively and the convolution integral can be written schematically as,

$$P_B(\vec{X}^R) = \int P_B(\vec{X}^G) T(\vec{X}^R | \vec{X}^G) d\vec{X}^G. \quad (5)$$

The set of variables  $\vec{X} \equiv (\vec{p}_T, Y, \phi, \hat{s}, M_1, M_2, \vec{\Omega})$  exhausts the twelve degrees of freedom (note that  $\vec{p}_T$  has 2 components and  $\vec{\Omega}$  contains 5 angles) available to the four

(massless) final state leptons. The differential volume element is given by  $d\vec{X} = d\hat{s}dM_1^2dM_2^2d\vec{\Omega} \cdot d\vec{p}_T dY d\phi$ .

To perform this convolution with the transfer function we must first transform to the basis in which the detector smearing of the lepton momenta is parameterized. This requires transforming from the basis of the twelve center of mass variables defined in Sec. III A to the three momentum basis for the four final state leptons. In this basis the lepton three momenta  $\vec{p}_i$  can be decomposed in terms of the component of the lepton momentum parallel to the direction ( $p_{i||}$ ) of motion and the two components perpendicular to the direction of motion ( $\vec{p}_{i\perp}$ ) (which are zero at generator level). We then make the assumption that detector smearing will only affect parallel components  $p_{i||}$  while the perpendicular components  $\vec{p}_{i\perp}$  are left invariant. Note that this assumption is equivalent to assuming angular resolution effects due to detector smearing can be neglected, which is an excellent approximation for the LHC detectors [53–55]. In the ( $p_{i||}, \vec{p}_{i\perp}$ ) basis only the transfer function associated with  $p_{i||}$  is non-trivial while the one associated with the perpendicular components can be represented simply as a delta function for each perpendicular direction, thus allowing for trivial integration over the eight  $\vec{p}_{i\perp}$  variables.

With these assumptions the integral in Eq.(5) can then be represented as follows,

$$P_B(\vec{X}^R) = \int P_B(\vec{X}^G) T(\vec{c} | \vec{P}^G) \times |\mathcal{J}_B| dc_1 dc_3 dM_1^{G^2} dM_2^{G^2}, \quad (6)$$

where we have defined the lepton momenta ‘smearing factors’  $c_i = p_{i||}^R / p_{i||}^G$  and,

$$\vec{c} = (c_1, c_2, c_3, c_4), \quad \vec{P}^G = (\vec{p}_1^G, \vec{p}_2^G, \vec{p}_3^G, \vec{p}_4^G). \quad (7)$$

We have also defined  $|\mathcal{J}_B|$  which is the  $12 \times 12$  Jacobian which parametrizes the (non-linear) transformation that takes us from the center of mass basis to the lepton smearing basis. The construction of this Jacobian is highly non trivial and requires a combination of analytic and numerical techniques which are beyond the scope of this overview, but the relevant details can be found in [40, 41].

We thus see in Eq.(6) that what started out as a twelve dimensional integral has been reduced to a much more manageable integration over four variables. The details and validation of this four dimensional integration, which is done using a recursive numerical integration technique [44] can also be found in [40, 41].

To construct the detector level signal differential cross section (again un-normalized), which is now a function of the effective couplings  $\vec{A}$ , we follow the same procedure as for the background starting from,

$$P_S(\vec{X}^R | \vec{A}) = \int P_S(\vec{X}^G | \vec{A}) T(\vec{X}^R | \vec{X}^G) d\vec{X}^G. \quad (8)$$

We again use the assumptions which allow us to perform the trivial integration over the eight  $\vec{p}_{i\perp}$  variables, but

instead transform to the following integration basis

$$P_S(\vec{X}^R | \vec{A}) = \int P_S(\vec{X}^G | \vec{A}) T(\vec{c} | \vec{P}^G) \times |\mathcal{J}_S| dc_1 d\hat{s}^G dM_1^{2G} dM_2^{2G} \quad (9)$$

We now also use the fact that, as mentioned below Eq.(4), the  $\hat{s}$  spectrum for the signal is  $\propto \delta(\hat{s}^G - m_h^2)$  (where  $m_h$  is the *generated* Higgs mass), enabling us to perform the integration over  $d\hat{s}^G$  as well. Thus, we have for the final signal detector level differential cross section,

$$P_S(\vec{X}^R | \vec{A}) = \int P_S(\vec{X}^G | \vec{A}) T(\vec{c} | \vec{P}^G) \times |\mathcal{J}_S| dc_1 dM_1^{2G} dM_2^{2G} \Big|_{\hat{s}^G = m_h^2}, \quad (10)$$

where again  $|\mathcal{J}_S|$  represent the  $12 \times 12$  Jacobian (which is different from  $|\mathcal{J}_B|$ ) taking us from the CM basis to the lepton smearing basis. By using a delta function to model the width of the resonance, there is one less dimension to integrate over as compared to the background case. While this makes it easier computationally in one respect, an additional complication arises since we have to integrate along a trajectory in which  $\hat{s}^G$  is kept constant. This places an additional constraint when performing the  $M_1^{G^2}, M_2^{G^2}$  integration which further complicates matters and must be properly taken into account. Explicit details of this integration and its validation along with the derivation of the signal Jacobian  $|\mathcal{J}_S|$  in Eq.(10) are given in [40, 41].

## B. Comments on Transfer Function

Detector response effects including effects from selection inefficiency may be parameterized into transfer functions in the following way,

$$T(c_i | \vec{p}_i^G) = \delta(\vec{p}_{i\perp}^R - \vec{p}_{i\perp}^G) \mathcal{S}(p_{i||}^R; p_{i||}^G) \times \epsilon(\vec{p}_{i\perp}^R, p_{i||}^R), \quad (11)$$

The response function  $\mathcal{S}$ , parameterizes the probability for a lepton with actual momentum  $\vec{p}_i^G$  to be reconstructed with momentum  $\vec{p}_i^R$ , while  $\delta$  is the Dirac delta function in the perpendicular components, and  $\epsilon$  is the selection efficiency. With typical lepton selection criteria employed by the LHC experiments [54, 55], it is a good approximation that each lepton is independent. Thus, the full transfer function for the event may be written as:

$$T(\vec{c} | \vec{P}^G) = \prod_{i=1}^4 T(c_i | \vec{p}_i^G). \quad (12)$$

We treat  $T(\vec{c} | \vec{P}^G)$  as a function of  $\vec{c}$  which takes the generator level momenta  $\vec{P}^G$  as input. The only effect of imperfect momentum measurement on the production spectra is to provide a small smearing of the  $\vec{p}_T$  spectrum for the four lepton system. We can mitigate the effects



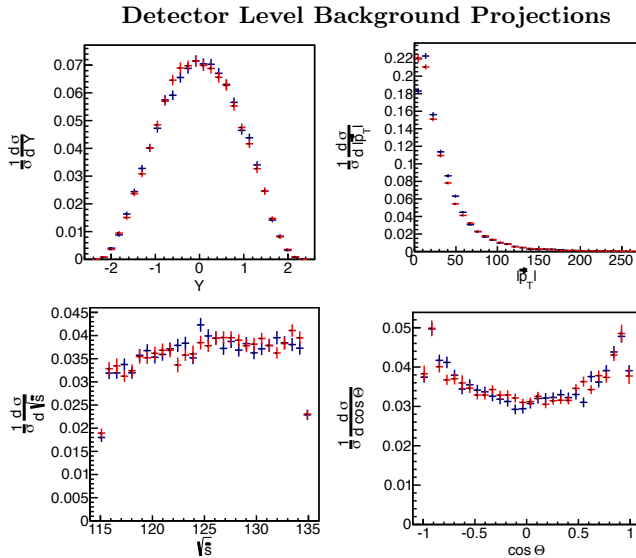


FIG. 2. Projections of the  $Y$ ,  $|\vec{p}_T|$ ,  $\sqrt{s} \equiv M_{4\ell}$  and  $\cos \Theta$  (see Sec. III A for definitions) spectra showing validation of the convolution described in Eq.(5) for the background. In *blue* we show the ‘boosted’ Madgraph sample with acceptance cuts and detector smearing applied while in *red* we show projections from our differential cross section after the convolution integration.

of the smearing by averaging over the production spectra when performing parameter extractions. Further details on the construction and implementation of the transfer function can be found in [38, 39, 41].

### C. Validation of Convolution Integral

As validation of the convolution integral we first show in Fig. 2-Fig. 5 projections for signal and background. We compare in these plots the distributions for a Madgraph sample which has had detector smearing and acceptance effects applied to it versus projections generated from our detector level differential cross sections obtained after the convolution described above.

We have obtained the signal and background production spectrum for the  $(\hat{s}, \vec{p}_T, Y, \phi)$  variables from POWHEG and boosted the Madgraph events and those from our projections accordingly. We have used the interpolation procedure described in Sec. III E to build the production spectra for the signal and background differential cross sections and combined them with the analytic expressions for the  $h \rightarrow 4\ell$  and  $q\bar{q} \rightarrow 4\ell$  processes. For the signal we show the *tree level* Standard Model point where  $\mathcal{A}_1^{ZZ} = 2$  and all other couplings are set to zero. For both signal and background we show only the  $2e2\mu$  final state, but results for  $4e$  (or  $4\mu$ ) are found in [38, 39, 41].

A further validation beyond these projections however is to look at the likelihoods (the differential cross section evaluated for a set of observables) for both the signal and

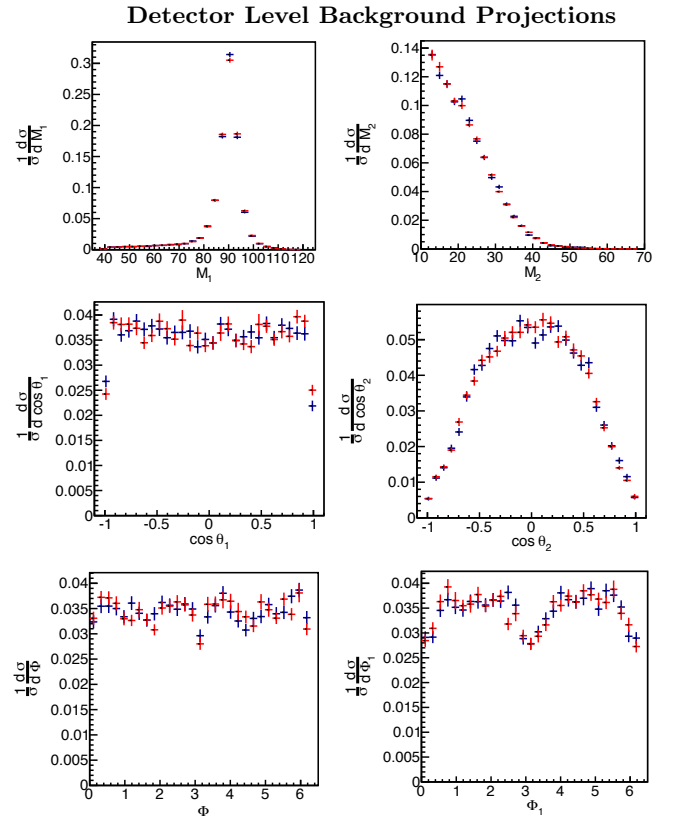


FIG. 3. Projections of the  $M_1$ ,  $M_2$ ,  $\cos \theta_1$ ,  $\cos \theta_2$ ,  $\Phi$  and  $\Phi_1$  (see Sec. III A for definitions) spectra showing validation of the convolution described in Eq.(5) for the background. In *blue* we show the ‘boosted’ Madgraph sample with acceptance cuts and detector smearing applied while in *red* we show projections from our differential cross section after the convolution integration.

background which contain the full correlations between the different variables. We show these in Fig. 6 for a CMS-like phase space and a very large number of events. To obtain these likelihoods we have evaluated our detector-level differential cross section with the Madgraph sample which has had detector smearing and acceptance effects applied and plotted it on top of the result of evaluating our detector-level differential cross section with events generated from the expression itself. We find the agreement between the two results to be very good. Further details are found in the accompanying documents [40, 41].

These plots should not be taken as validation of the complete detector-level differential cross sections which must be validated with full simulation and data. They are meant only to show the validation of the convolution procedure as well as the construction of the generator-level differential cross sections including the analytic computations. Complete validations of the full detector level likelihoods including the various production and background effects can be found in [38, 39, 41].

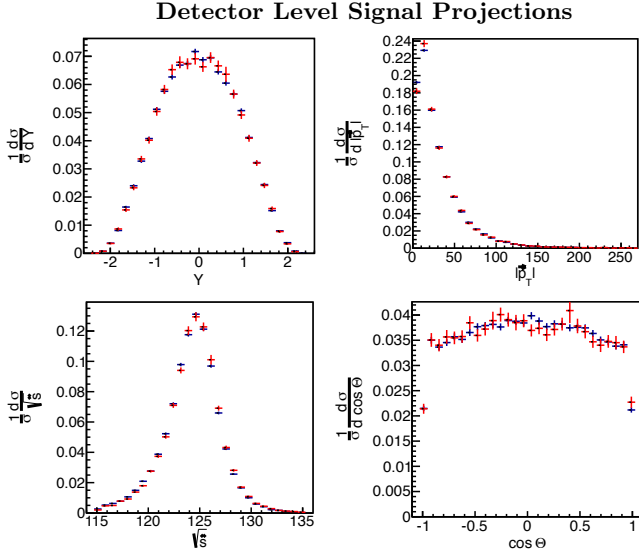


FIG. 4. Projections of the  $Y$ ,  $|\vec{p}_T|$ ,  $\sqrt{\hat{s}} \equiv M_{4\ell}$  and  $\cos \Theta$  (see Sec. III A for definitions) spectra showing validation of the convolution described in Eq.(8) for the *tree level* SM signal. In *blue* we show the ‘boosted’ Madgraph sample with acceptance cuts and detector smearing applied while in *red* we show projections from our differential cross section after the convolution integration.

## V. CONSTRUCTION OF LIKELIHOODS AND PARAMETER EXTRACTION

With the detector level differential cross sections obtained in Eq.(6) and Eq.(10) in hand we can then go on to construct the full likelihood for a particular dataset. Before doing so, we must properly normalize the background and signal differential cross sections by performing the full integration over all twelve *reconstructed*  $\vec{X}$  variables where from now on we drop the superscript  $R$  since we only deal with detector level observables in what follows. In this section we present a schematic overview of the normalization procedure. We also at this stage briefly discuss averaging over the production variables ( $Y, \vec{p}_T, \phi$ ) and the implementation of systematic uncertainties through the use of nuisance parameters in the likelihood functions. Further details can be found in [40, 41].

### A. Normalization of Background and Signal

One can reduce the effects of production uncertainties by averaging over the detector level production variables ( $Y, \vec{p}_T, \phi$ ). This is straightforwardly done for the background differential cross sections by the following 4-dimensional integration,

$$P_B(\hat{s}, M_1, M_2, \vec{\Omega}) = \int P_B(\vec{X}) dY d\vec{p}_T d\phi. \quad (13)$$

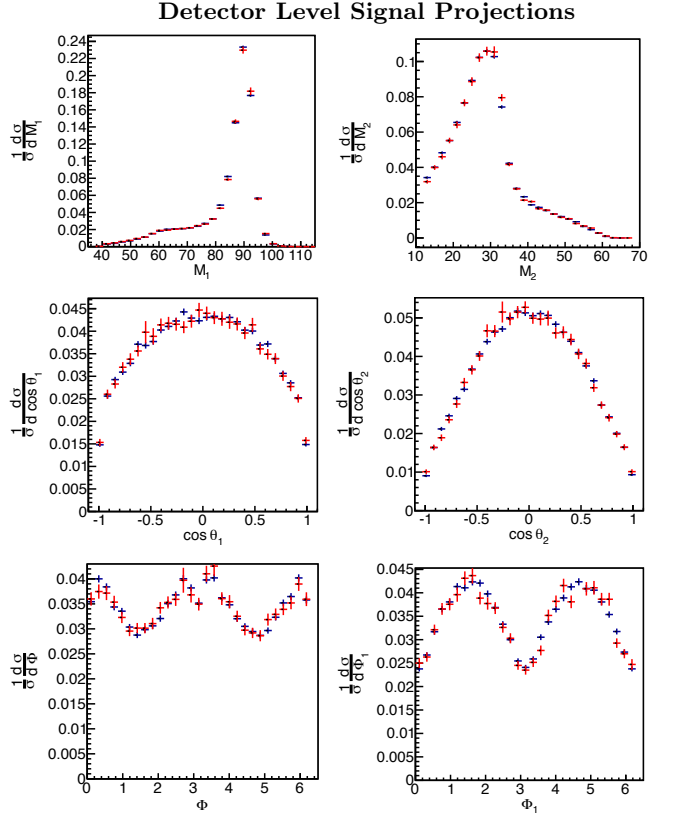


FIG. 5. Projections of the  $M_1$ ,  $M_2$ ,  $\cos \theta_1$ ,  $\cos \theta_2$ ,  $\Phi$  and  $\Phi_1$  (see Sec. III A for definitions) spectra showing validation of the convolution described in Eq.(8) for the *tree level* SM signal. In *blue* we show the ‘boosted’ Madgraph sample with acceptance cuts and detector smearing applied while in *red* we show projections from our differential cross section after the convolution integration.

An overall volume factor is not shown because for the purpose of likelihood maximization this constant factor is not relevant. What matters is that the relative normalization between all components in the likelihood is done consistently. With this differential cross sections in terms of the eight center of mass decay observables we can obtain the overall normalization via a Monte Carlo integration procedure described in [40, 41],

$$\mathcal{N}_B = \int P_B(\hat{s}, M_1, M_2, \vec{\Omega}) \times d\hat{s} dM_1^2 dM_2^2 d\vec{\Omega}, \quad (14)$$

which gives our final normalized background *pdf* as,

$$P_B(\hat{s}, M_1, M_2, \vec{\Omega}) = \mathcal{N}_B^{-1} \times P_B(\hat{s}, M_1, M_2, \vec{\Omega}). \quad (15)$$

We have calculated the  $q\bar{q} \rightarrow 4\ell$  expression as a sum of the separate individual contributions [19, 35] making it possible to easily perform the integration on each smaller piece to obtain each normalization and then simply sum over them to obtain the overall normalization.

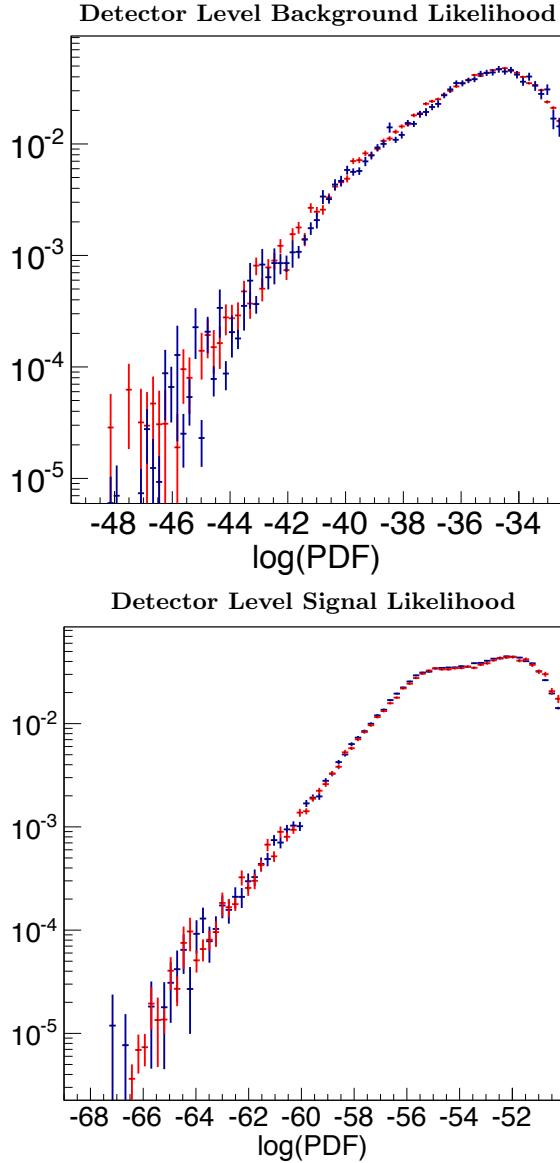


FIG. 6. Validation of the convolution integrals described in Eq.(5) and Eq.(8). In *blue* we show the ‘boosted’ Madgraph sample with acceptance cuts and detector smearing applied while in *red* we show projections from our differential cross sections after the convolution integration for the *tree level* SM signal and background likelihood.

Similarly for the signal we have for the averaging over  $(Y, \vec{p}_T, \phi)$  variables,

$$P_S(\hat{s}, M_1, M_2, \vec{\Omega}|\vec{\mathcal{A}}) = \int P_S(\vec{X}|\vec{\mathcal{A}}) dY d\vec{p}_T. \quad (16)$$

To obtain the overall normalization in the signal case we first note that it is a function of the underlying parameters  $\vec{\mathcal{A}}$  defined in Eq.(2)). However, from the calculation of the parton level differential cross section presented in [19, 35] or from considering Eq.(2) it is clear (assuming constant effective couplings) that  $P_S(\hat{s}, M_1, M_2, \vec{\Omega}|\vec{\mathcal{A}})$

is a sum over terms each of which is proportional to  $\mathcal{A}_n^i \mathcal{A}_m^{j*}$ . Thus we can write,

$$P_S(\hat{s}, M_1, M_2, \vec{\Omega}|\vec{\mathcal{A}}) = \sum_{ij} \sum_{nm} \mathcal{A}_n^i \mathcal{A}_m^{j*} \times P_S(\hat{s}, M_1, M_2, \vec{\Omega})_{nm}^{ij}, \quad (17)$$

where  $P_S(\hat{s}, M_1, M_2, \vec{\Omega})_{nm}^{ij}$  represents the individual differential cross sections with the couplings factored out. The separate normalizations for each term can now easily be obtained via,

$$\mathcal{N}_{nm}^{ij} = \int P_S(\hat{s}, M_1, M_2, \vec{\Omega})_{nm}^{ij} \times d\hat{s} dM_1^2 dM_2^2 d\vec{\Omega}, \quad (18)$$

from which we can now obtain the total overall normalization for the signal *pdf* as,

$$\mathcal{N}_S(\vec{\mathcal{A}}) = \sum_{ij} \sum_{nm} \mathcal{A}_n^i \mathcal{A}_m^{j*} \times \mathcal{N}_{nm}^{ij}. \quad (19)$$

This gives finally for the normalized signal *pdf*,

$$\mathcal{P}_S(\hat{s}, M_1, M_2, \vec{\Omega}|\vec{\mathcal{A}}) = \mathcal{N}_S^{-1}(\vec{\mathcal{A}}) \times P_S(\hat{s}, M_1, M_2, \vec{\Omega}|\vec{\mathcal{A}}). \quad (20)$$

Since each  $\mathcal{N}_{nm}^{ij}$  is computed, one does not need to compute the normalization each time a new hypothesis for  $\vec{\mathcal{A}}$  is constructed. The procedure outlined here also works on more general polynomial functions of the parameters  $\vec{\mathcal{A}}$  which one finds after expanding potentially momentum-dependent form factors in powers of momenta. See [40] for this more general discussion.

Note also that if we take the Higgs mass as a fixed input and only fit for ratios of parameters and not their overall normalization, we do not need the absolute normalization of the differential cross sections. It thus suffices to have the relative normalization between the different components correct when performing the maximization. This fact greatly reduces the computational complexity. Instead of propagating the full normalization and aligning units correctly so that when one integrates over all 8 dimensions unity is obtained, it is sufficient to do a Monte Carlo integration using a fixed sample size in a consistent and sufficiently large range. The meaning of the log likelihood difference remains unchanged with this construction. Further details of the normalization procedure for both signal and background are found in [40, 41].

## B. Signal Plus Background *pdf* and Final Likelihood

With Eq.(15) and Eq.(20) in hand we can now build the signal plus background *pdf* from which the total likelihood will be constructed. The signal plus background

*pdf* can be written as,

$$\begin{aligned} \mathcal{P}_{S+B}(\mathcal{O}|F_B^i, \vec{\mathcal{A}}) &= \sum_i F_B^i \times \mathcal{P}_B^i(\hat{s}, M_1, M_2, \vec{\Omega}) \\ &+ (1 - \sum_i F_B^i) \times \mathcal{P}_S(\hat{s}, M_1, M_2, \vec{\Omega}|\vec{\mathcal{A}}). \end{aligned} \quad (21)$$

where  $\mathcal{O} \equiv (\hat{s}, M_1, M_2, \vec{\Omega})$  is our final set of observables to be used in the construction of the likelihood and  $F_B^i$  is the background fraction for a particular component, each of which must also be extracted.

The sum over background components is given by,

$$\begin{aligned} \sum_i F_B^i \times \mathcal{P}_B^i(\mathcal{O}) &= F_B^{q\bar{q}} P_B^{q\bar{q}}(\mathcal{O}) \\ &+ F_B^{gg} P_B^{gg}(\mathcal{O}) + F_B^{Z+X} P_B^{Z+X}(\mathcal{O}), \end{aligned} \quad (22)$$

where  $P_B^{q\bar{q}}(\mathcal{O})$  is the dominant  $q\bar{q} \rightarrow 4\ell$  component and is obtained via the convolution integral in Eq.(1). The sub-dominant  $gg \rightarrow 4\ell$  and  $Z + X$  components, given by  $P_B^{gg}(\mathcal{O})$  and  $P_B^{Z+X}(\mathcal{O})$  respectively, must be obtained via the linearly interpolated ‘look-up’ tables from large Monte Carlo samples as discussed in Sec. III F.

We can now write the likelihood of obtaining a particular dataset containing  $N$  events as,

$$L(F_B^i, \vec{\mathcal{A}}) = \prod_{\mathcal{O}}^N \mathcal{P}_{S+B}(\mathcal{O}|F_B^i, \vec{\mathcal{A}}). \quad (23)$$

This likelihood can also be combined with an appropriate poisson weighting factor to account for the probability of observing a given number of events [38, 39, 41]. In the case of multiple final states (for example  $4e$ ,  $4\mu$  and  $2e2\mu$ ), we build the likelihood function and implement the appropriate systematic uncertainties for each one separately. We now briefly discuss the implementation of the systematic uncertainties.

### C. Including Systematic Uncertainties

Systematic uncertainties must be accounted for given our imperfect knowledge of various aspects of the analysis procedure. The lepton momentum resolution, the size of the backgrounds, and the exact production spectra are some important examples. For each of these systematic uncertainties we can associate an undetermined parameter which parametrizes our ignorance of the corresponding effect. Since we are not directly interested in these parameters, but only use them to estimate our systematic uncertainties, they are deemed nuisance parameters and are subsequently profiled over [38, 39].

This is done by generating alternative *pdfs* using different values for the nuisance parameter of interest. To give one important example, we generate *pdfs* with narrower or wider lepton response functions to parameterize our knowledge of the lepton momentum resolution. If we define the nominal *pdf* to be  $\mathcal{P}_0(\mathcal{O})$  and the alternative as  $\mathcal{P}_1(\mathcal{O})$ , one can parameterize the dependence of the

likelihood on a nuisance parameter  $n$  by interpolating between the nominal and the alternative *pdfs* as follows:

$$\begin{aligned} \mathcal{P}(\mathcal{O}|n) &= (1 - n) \mathcal{P}_0(\mathcal{O}) + n \mathcal{P}_1(\mathcal{O}) \\ &= P_0(\mathcal{O}) + n [P_1(\mathcal{O}) - P_0(\mathcal{O})]. \end{aligned} \quad (24)$$

It is instructive to observe that, for all values of  $n$ , the normalization of the total *pdf* stays the same. Given the asymmetric nature of many systematic uncertainties, it is more appropriate to generate many ‘check-points’ along the axis of  $n$  and to do piece-wise interpolation without the need of worrying about the normalization. Non-central values of  $n$  are *a priori* disfavored, therefore one can impose a prior on top of the interpolated likelihood:

$$\mathcal{P}(\mathcal{O}|n) = \mathcal{P}(\mathcal{O}|n) \mathcal{G}(n), \quad (25)$$

where  $\mathcal{G}(n)$  is typically a Gaussian centered at the central value of  $n$ . In the case of multiple systematic uncertainties, one can replace  $n$  by a vector of nuisance parameters  $\vec{n}$ , and the prior  $\mathcal{G}(n)$  by  $\mathcal{G}(\vec{n})$ . In general  $\mathcal{G}(\vec{n})$  is a multivariate Gaussian-like function with primary axes which are some combination of different nuisance parameter directions. However one can carefully define the nuisance parameters such that correlations between them are negligible. In this limit  $\mathcal{G}(\vec{n})$  can be written as the product of many Gaussian-like functions. This procedure for including systematic uncertainties has been implemented in a recent CMS analysis utilizing our framework [38, 39] and further details can be found in [41].

### D. Comments on Parameter Extraction

As discussed in [13, 17, 32] the advantage of analytic approaches is that the likelihood can be maximized for a large set of parameters in the most optimal way without losing information. Our framework allows for the ‘analytic’ nature of these approaches to be maintained at detector level giving us the ability to perform fast and accurate multi-parameter fits for lagrangian parameters directly from the data. This is possible once the convolution in Eq.(1) is performed and after normalization of the signal and background *pdfs* allowing us to obtain the full detector level likelihood  $L(\vec{\mathcal{A}})$  for a particular dataset. With the likelihood in hand a maximization procedure to find the *global* maximum can be performed to obtain the value of the parameters for which the likelihood is maximized. For this task we have incorporated the well established MINUIT [63] function minimization/maximization code into our framework. We find excellent rates of convergence and a high degree of stability in locating the global maximum of the likelihood as well as accurate extraction of the parameters as demonstrated in [38–41] where more details can be found.

One important feature of the procedure is that the computationally intensive component of evaluating the likelihood only needs to be done for the events in the final dataset used in the fit for a given experiment. Therefore

the computationally expensive pieces can be calculated on the computing grid *prior to the analysis of the data*, and the fit for parameter extraction itself is then completed within a few seconds. This allows for a great deal of flexibility, including testing alternative parameterizations, when fitting the undetermined parameters. Many examples of the types of parameter extractions which can be done within our framework, both at generator and at detector level, can be found in [35, 37–39, 41].

## VI. SUMMARY AND CONCLUSIONS

In this study we build upon an earlier study [35] to construct a comprehensive analysis framework aimed at extracting as much information as possible from the Higgs golden channel. Our framework is based on a maximum likelihood method constructed from analytic expressions of the fully differential cross sections for the  $h \rightarrow 4\ell$  decay as well as the dominant irreducible  $q\bar{q} \rightarrow 4\ell$  background which were computed in [19, 35]. As our main result, we have constructed the full 12-dimensional detector level likelihood utilizing all observables available in the golden channel. This allows us to perform parameter extraction of the various possible Higgs couplings, including general CP odd/even admixtures and any possible phases.

The detector-level likelihood is obtained by the explicit convolution of a transfer function, encapsulating the relevant detector effects, with the generator-level probability density formed out of the signal and background differential cross sections. After performing this 12-dimensional convolution integral and its normalization we obtain a probability density function from which we construct an un-binned detector-level likelihood which is a continuous function of the effective couplings.

In summary we have given broad overview of a framework optimized for extracting Higgs couplings in the golden channel. We have sketched how the convolution is performed and shown various validations as well as discussed the principles and theoretical basis of the framework. Many of the technical details as well as results using our framework can be found in [19, 35, 37, 38, 40, 41]. This framework has already proved useful in a recent CMS analysis [38, 39] and can be used in the future in a variety of ways to study Higgs couplings in the golden channel using data obtained at the LHC and other future colliders.

**Acknowledgments:** The authors are grateful to Artur Apresyan, Michalis Bachtis, Adam Falkowski, Patrick Fox, Andrei Gritsan, Roni Harnik, Alex Mott, Pavel Nadolsky, Daniel Stolarski, Reisaburo Tanaka, Nhan Tran, Roberto Vega, and Felix Yu for helpful discussions as well as the ATLAS and CMS collaborations for their encouragement and interest in this work. R.V.M. is supported by the Fermilab Graduate Student Fellowship in Theoretical Physics and the ERC Advanced Grant Higgs@LHC. Fermilab is operated by Fermi Research Alliance, LLC, under Contract No. DE-AC02-07CH11359 with the United States Department of Energy. Y.C. is supported by the Weston Havens Foundation and DOE grant DE-FG02-92-ER-40701. This work is also sponsored in part by the DOE grant No. DE-FG02-91ER40684. R.V.M. is also grateful to the CalTech physics department for their hospitality during which much of this work was done. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number OCI-1053575.

- 
- [1] G. Aad et al. (ATLAS Collaboration), Phys.Lett. **B716**, 1 (2012), 1207.7214.
  - [2] S. Chatrchyan et al. (CMS Collaboration), Phys.Lett. **B716**, 30 (2012), 1207.7235.
  - [3] C. A. Nelson, Phys.Rev. **D37**, 1220 (1988).
  - [4] A. Soni and R. Xu, Phys.Rev. **D48**, 5259 (1993), hep-ph/9301225.
  - [5] D. Chang, W.-Y. Keung, and I. Phillips, Phys.Rev. **D48**, 3225 (1993), hep-ph/9303226.
  - [6] V. D. Barger, K.-m. Cheung, A. Djouadi, B. A. Kniehl, and P. Zerwas, Phys.Rev. **D49**, 79 (1994), hep-ph/9306270.
  - [7] T. Arens and L. Sehgal, Z.Phys. **C66**, 89 (1995), hep-ph/9409396.
  - [8] S. Choi, . Miller, D.J., M. Muhlleitner, and P. Zerwas, Phys.Lett. **B553**, 61 (2003), hep-ph/0210077.
  - [9] C. Buszello, I. Fleck, P. Marquard, and J. van der Bij, Eur.Phys.J. **C32**, 209 (2004), hep-ph/0212396.
  - [10] R. M. Godbole, . Miller, D.J., and M. M. Muhlleitner, JHEP **0712**, 031 (2007), 0708.0458.
  - [11] V. Kovalchuk, J.Exp.Theor.Phys. **107**, 774 (2008).
  - [12] Q.-H. Cao, C. Jackson, W.-Y. Keung, I. Low, and J. Shu, Phys.Rev. **D81**, 015010 (2010), 0911.3398.
  - [13] Y. Gao, A. V. Gritsan, Z. Guo, K. Melnikov, M. Schulze, et al., Phys.Rev. **D81**, 075022 (2010), 1001.3396.
  - [14] A. De Rujula, J. Lykken, M. Pierini, C. Rogan, and M. Spiropulu, Phys.Rev. **D82**, 013003 (2010), 1001.5300.
  - [15] J. S. Gainer, K. Kumar, I. Low, and R. Vega-Morales, JHEP **1111**, 027 (2011), 1108.2274.
  - [16] B. Coleppa, K. Kumar, and H. E. Logan (2012), 1208.2692.
  - [17] S. Bolognesi, Y. Gao, A. V. Gritsan, K. Melnikov, M. Schulze, et al. (2012), 1208.4018.
  - [18] D. Stolarski and R. Vega-Morales, Phys.Rev. **D86**, 117504 (2012), 1208.4840.
  - [19] Y. Chen, N. Tran, and R. Vega-Morales, JHEP **1301**, 182 (2013), 1211.1959.
  - [20] R. Boughezal, T. J. LeCompte, and F. Petriello (2012), 1208.4311.
  - [21] A. Belyaev, N. D. Christensen, and A. Pukhov (2012), 1207.6082.
  - [22] P. Avery, D. Bourilkov, M. Chen, T. Cheng, A. Drozd-

- skiy, et al. (2012), 1210.0896.
- [23] J. M. Campbell, W. T. Giele, and C. Williams (2012), 1205.3434.
  - [24] J. M. Campbell, W. T. Giele, and C. Williams (2012), 1204.4424.
  - [25] S. Chatrchyan et al. (CMS Collaboration), JHEP **1204**, 036 (2012), 1202.1416.
  - [26] S. Chatrchyan et al. (CMS Collaboration), Phys.Lett. **B716**, 30 (2012), 1207.7235.
  - [27] S. Chatrchyan et al. (CMS Collaboration), Phys.Rev.Lett. **110**, 081803 (2013), 1212.6639.
  - [28] T. Modak, D. Sahoo, R. Sinha, and H.-Y. Cheng (2013), 1301.5404.
  - [29] Y. Sun, X.-F. Wang, and D.-N. Gao (2013), 1309.4171.
  - [30] J. S. Gainer, J. Lykken, K. T. Matchev, S. Mrenna, and M. Park, Phys.Rev.Lett. **111**, 041801 (2013), 1304.4936.
  - [31] P. Artoisenet, P. de Aquino, F. Demartin, R. Frederix, S. Frixione, et al. (2013), 1306.6464.
  - [32] I. Anderson, S. Bolognesi, F. Caola, Y. Gao, A. V. Gritsan, et al. (2013), 1309.4819.
  - [33] M. Chen, T. Cheng, J. S. Gainer, A. Korytov, K. T. Matchev, et al. (2013), 1310.1397.
  - [34] G. Buchalla, O. Cata, and G. D'Ambrosio (2013), 1310.2574.
  - [35] Y. Chen and R. Vega-Morales, JHEP **1404**, 057 (2014), 1310.2893.
  - [36] J. S. Gainer, J. Lykken, K. T. Matchev, S. Mrenna, and M. Park (2014), 1403.4951.
  - [37] Y. Chen, R. Harnik, and R. Vega-Morales, Phys.Rev.Lett. **113**, 191801 (2014), 1404.1336.
  - [38] Tech. Rep. CMS-PAS-HIG-14-014, CERN, Geneva (2014), CMS Collaboration.
  - [39] V. Khachatryan et al. (CMS Collaboration) (2014), 1411.3441.
  - [40] Y. Chen, E. Di Marco, J. Lykken, M. Spiropulu, R. Vega-Morales, et al. (2014), 1410.4817.
  - [41] Y. Chen (2014), Precision measurement of the 125 GeV Higgs boson discovered in proton-proton collisions at  $\sqrt{s} = 7, 8$  TeV with the CMS detector at the LHC, Thesis in preparation for P.h.D. at California Institute of Technology.
  - [42] Y. Gao (2013), Talk given at ATLAS Higgs WG Workshop, <https://indico.cern.ch/event/280998/session/5/contribution/29/material/slides/0.pdf>.
  - [43] R. Vega-Morales (2013), The Higgs Boson as a Window to Beyond the Standard Model, Thesis in fulfillment of P.h.D requirements, Northwestern University.
  - [44] B. Bradie, *A friendly introduction to numerical analysis: with C and MATLAB materials on website; int. ed.* (Prentice-Hall, Englewood Cliffs, NJ, 2006).
  - [45] P. Artoisenet and O. Mattelaer, PoS **CHARGED2008**, 025 (2008).
  - [46] P. Artoisenet, V. Lemaitre, F. Maltoni, and O. Mattelaer, JHEP **1012**, 068 (2010), 1007.3300.
  - [47] P. Artoisenet, P. de Aquino, F. Maltoni, and O. Mattelaer (2013), 1304.6414.
  - [48] A. Falkowski and R. Vega-Morales (2014), 1405.1095.
  - [49] A. Bredenstein, A. Denner, S. Dittmaier, and M. Weber, Phys.Rev. **D74**, 013004 (2006), hep-ph/0604011.
  - [50] A. Bredenstein, A. Denner, S. Dittmaier, and M. Weber, Nucl.Phys.Proc.Suppl. **160**, 131 (2006), hep-ph/0607060.
  - [51] N. Adam, T. Aziz, J. Andersen, A. Belyaev, T. Binoth, et al., pp. 215–289 (2008), 0803.1154.
  - [52] J. Neyman and E. S. Pearson (1933), Philosophical Transactions of the Royal Society of London. Series A Vol. 231, pp. 289–337.
  - [53] CMS (2006), CMS Physics: Technical Design Report Volume 1: Detector Performance and Software.
  - [54] CMS (2013), <https://twiki.cern.ch/twiki/bin/view/CMS/Public/EGMElectronsMoriand2013>.
  - [55] CMS (2013), Electron performance with 19.6 fb<sup>-1</sup> of data collected at  $\sqrt{s} = 8$  TeV with the CMS detector.
  - [56] Y. Chen and R. Vega-Morales (2013), Website under construction: <http://yichen.me/project/GoldenChannel/>.
  - [57] N. D. Christensen and C. Duhr, Comput.Phys.Commun. **180**, 1614 (2009), 0806.4194.
  - [58] J. Alwall, P. Demin, S. de Visscher, R. Frederix, M. Herquet, et al., JHEP **0709**, 028 (2007), 0706.2334.
  - [59] Y. Chen, D. Stolarski, and R. Vega-Morales (2014), Work in preparation.
  - [60] A. Falkowski, F. Riva, and A. Urbano (2013), 1303.1812.
  - [61] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, JHEP **1106**, 128 (2011), 1106.0522.
  - [62] T. Melia, P. Nason, R. Rontsch, and G. Zanderighi, JHEP **1111**, 078 (2011), 1107.5051.
  - [63] F. James (1994), MINUIT Function Minimization and Error Analysis: Reference Manual Version 94.1.